Project 1: Regression: predicting compressive strength of concrete

Problem Statement: build a regression model that predicts compressive strength from inputs (cement, blast furnace slag, fly ash, water, aggregates, age). Success criteria: RMSE < target and $R^2 > 0.7$ on test set.

Dataset:

Main source: https://archive.ics.uci.edu/dataset/165/concrete+compressive+strength

Kaggle source: https://www.kaggle.com/datasets/elikplim/concrete-compressive-strength-data-set

Dataset size: 1030 rows, 9 columns.

ML Algorithms used:

1. Baseline linear regression,
2. Random forest regression,
3. K-fold cross validation.

Results/Post Processing:

1. Linear Regresion: RMSE, R^2. Predicated vs actual plot.
2. Random Forest Regression: Feature importance, residual distribution.